

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

APPLICATION FOR LETTERS PATENT

**Transport System for Instant Messaging**

Inventor(s):

John Holmes

David Michael Miller

Walter vonKoch

ATTORNEY'S DOCKET NO. MS1-1526US

CLIENT'S DOCKET NO. 303261.1

**EV317722487**

1                                   **TRANSPORT SYSTEM FOR INSTANT MESSAGING**

2                                   **CROSS-REFERENCE TO RELATED APPLICATION**

3                   The present application is related to co-pending U.S. patent application Ser. No.  
4                   \_\_\_\_\_, Attorney Docket No. MS1-1527, entitled "Instant Messaging Object  
5                   Store," by David Michael Miller, John Holmes, and Walter vonKoch, which is filed  
6                   concurrently herewith, assigned to the assignee of the present application, and  
7                   incorporated herein by reference for all that it teaches and discloses.  
8

9  
10                                   **TECHNICAL FIELD**

11                   The described subject matter relates to computer communications. More  
12                   particularly, the subject matter relates to a transport system for instant messaging.  
13

14                                   **BACKGROUND**

15                   Instant messaging is becoming a very popular communications tool for users of  
16                   computer devices. An instant messaging (IM) application (e.g., WINDOWS ® Messenger  
17                   system of Microsoft Corporation of Redmond, WA, Yahoo! Messenger, AOL instant  
18                   messenger (AIM), and the like) enables a user to engage in a real time conversation with  
19                   one or more contacts, who are identified in the user's private list of contacts. Typically,  
20                   private lists are stored on a server and a conversation is established through a switchboard,  
21                   or relay server, which directs inbound messages to the appropriate recipients.  
22

23                   Routing real time messages through the switchboard can be slower and more costly  
24                   (in terms of server resources) than instant messaging in a peer-to-peer arrangement, wherein  
25

1 the messages do not go through a switchboard, but pass directly to and from participants in  
2 the conversation. Thus, after a conversation is established using a switchboard server, a  
3 typical instant messaging system may attempt to convert the conversation into a peer-to-  
4 peer configuration. Converting to peer-to-peer may be done a number of ways, typically  
5 involving identifying a peer by an internet protocol (IP) address and/or port number,  
6 whereby messages can be sent directly to the peer.

7 However, creating a peer-to-peer conversation can be difficult if not impossible in  
8 situations involving active network devices, such as firewalls or network address translators  
9 (NATs). Such devices are used for system security, but are typically designed to prevent  
10 direct access to computers behind such devices. For example, a NAT has its own IP  
11 address, and computers behind the NAT have their own IP addresses, which are private; i.e.,  
12 other computers outside the NAT may not be able to determine the IP addresses of the  
13 computers behind the NAT. The IP address of a message received by the NAT is translated  
14 to the IP address of the recipient computer behind the NAT. Various techniques may be  
15 applied to tunnel through or traverse an active network device to form a direct connection.  
16

17 As instant messaging systems improve, they typically provide more features, besides  
18 text, that make the conversation a richer experience. For example, an instant messaging  
19 system may allow a user to transmit audio, video, or a custom user tile (also referred to as a  
20 Display Picture) uniquely associated with the user. Such features typically require more  
21 resources, such as transmission bandwidth, than text to include in a conversation. Thus, a  
22 peer-to-peer connection is desirable when carrying on an instant messaging conversation  
23 because peer-to-peer typically provides higher performance than a connection through a  
24 switchboard server.  
25

## **SUMMARY**

Implementations described and claimed herein solve the discussed problems, and other problems.

An exemplary method involves creating a master session over a first connection through a server; and creating a virtual channel over the connection, the virtual channel operable to communicate a feature session. The method may involve establishing a direct connection that bypasses the server, and switching communication of the feature session to the direct connection.

A system for network communication includes a plurality of transport bridges, each transport bridge corresponding to an active network device configuration, and a switching module operable to choose one of the transport bridges to form a connection between two computing devices based on the active network device configuration.

## **BRIEF DESCRIPTION OF THE DRAWINGS**

Fig. 1 is a block diagram depicting components in an exemplary instant messaging system.

Fig. 2 illustrates a transport system including exemplary layers for managing communication of data in an instant messaging scenario.

Fig. 3 is a block diagram illustrating an exemplary bridge selection module operable to select a preferred transport bridge.

Fig. 4 is a block diagram illustrating virtual channels across a physical connection between two peer devices.

1 Fig. 5 illustrates an operation flow having exemplary operations for establishing a  
2 feature session within a master session and creating a direct connection if possible.

3 Fig. 6 illustrates an operation flow having exemplary operations for determining a  
4 preferred transport bridge and creating a direct connection using the preferred transport  
5 bridge.

6 Fig. 7 illustrates an operation flow having exemplary operations for transporting  
7 binary large objects (BLOBs) in sequential order.

8 Fig. 8 illustrates an exemplary system that provides a suitable operating  
9 environment to transparently select a preferred transport bridge and communicate feature  
10 data in an instant messaging scenario.

## 11 **DETAILED DESCRIPTION**

12  
13  
14 Turning to the drawings, wherein like reference numerals refer to like elements,  
15 various methods are illustrated as being implemented in a suitable computing  
16 environment. Although not required, various exemplary methods will be described in the  
17 general context of computer-executable instructions, such as program modules, being  
18 executed by a personal computer and/or other computing device. Generally, program  
19 modules include routines, programs, objects, components, data structures, etc. that  
20 perform particular tasks or implement particular abstract data types.

21 Moreover, those skilled in the art will appreciate that various exemplary methods  
22 may be practiced with other computer system configurations, including hand-held  
23 devices, multi-processor systems, microprocessor based or programmable consumer  
24 electronics, network PCs, minicomputers, mainframe computers, and the like. Various  
25

1 exemplary methods may also be practiced in distributed computing environments where  
2 tasks are performed by remote processing devices that are linked through a  
3 communications network. In a distributed computing environment, program modules  
4 may be located in both local and remote memory storage devices.

5 In some diagrams herein, various algorithmic acts are summarized in individual  
6 “blocks”. Such blocks describe specific actions or decisions that are made or carried out  
7 as a process proceeds. Where a microcontroller (or equivalent) is employed, the flow  
8 charts presented herein provide a basis for a “control program” or software/firmware that  
9 may be used by such a microcontroller (or equivalent) to effectuate the desired control.  
10 As such, the processes are implemented as machine-readable instructions storable in  
11 memory that, when executed by a processor, perform the various acts illustrated as  
12 blocks.

13 Those skilled in the art may readily write such a control program based on the  
14 flow charts and other descriptions presented herein. It is to be understood and  
15 appreciated that the subject matter described herein includes not only devices and/or  
16 systems when programmed to perform the acts described below, but the software that is  
17 configured to program the microcontrollers and, additionally, any and all computer-  
18 readable media on which such software might be embodied. Examples of such computer-  
19 readable media include, without limitation, floppy disks, hard disks, CDs, RAM, ROM,  
20 flash memory and the like.  
21  
22

### 23 **Exemplary Network Environment for Instant Messaging**

24  
25

1 Fig. 1 illustrates an exemplary network environment 100 for an instant messaging  
2 conversation. Generally, two clients 102 and 104 can communicate with each other via a  
3 network 106. Active network devices 108 and 110 may or may not be present in the  
4 network environment 100. An echo server 112 and a switchboard server 114 may be  
5 connected via the network 106 to facilitate communication between clients 102 and 104.

6 An instant messaging (IM) platform 116 enables the client (1) 102 and the client  
7 (2) 104 to engage in an instant messaging conversation. A user of the IM platform 116  
8 interacts with the IM platform 116 via a user interface (not shown) to send and receive  
9 messages to and from the client (2) 104. The IM platform 116 includes features 118 (also  
10 called end user features (EUFs)) and a transport protocol stack 122 that facilitate  
11 communication of data between the client (1) 102 and the client (2) 104.  
12

13 More specifically, the transport protocol stack 122 transparently establishes an  
14 instant messaging session based on characteristics of the network environment 100. In  
15 general, the established session may be through the switchboard (SB) server 114 or,  
16 alternatively, via a peer-to-peer connection 124, whereby data is not routed through the  
17 SB server 114.

18 The clients 102 and 104 may access the network 106 via an internet service  
19 provider (not shown), or otherwise. In one embodiment, the clients 102 and 104 each  
20 have unique network addresses, ports, and/or identifiers, whereby communications can be  
21 established between the clients 102 and 104. For example, the SB server 114 can route  
22 data to the client computer (2) 104 using the client computer's (2) 104 network address,  
23 port, and/or identifier. Over a peer-to-peer connection 124, the client computer (1) 102  
24  
25



1 can directly send data to the client computer (2) 104, using the client computer's (2) 104  
2 network address, port, and/or identifier.

3 As discussed, active network device (1) 106 and active network device (2) 110  
4 may or may not be present. Examples of active network devices are firewalls and  
5 network address translators, which provide system security for their respective client  
6 devices. Thus, the active network device (2) 110 may provide security for the client  
7 computer (2) 104, and the active network device (1) 108 may provide security for the  
8 client computer (2) 102.

9 Various network configurations are therefore possible. In one network  
10 configuration, the active network device (1) 108 is present, but the active network device  
11 (2) 110 is not present. In another network configuration, the active network device (2)  
12 110 is present, but the active network device (1) 108 is not present. In another network  
13 configuration, neither the active network device (1) 108, nor the active network device (2)  
14 110 are present. In a fourth network configuration, both the active network device (1)  
15 108 and the active network device (2) 110 are present.

17 The active network devices 108 and 110 may have unique network addresses,  
18 such as internet protocol (IP) addresses, whereby the active network devices 108 and 110  
19 receive data bound for their respective clients 102 and 104. Active network devices 108  
20 or 110 receive data on their network addresses and translate the network addresses to the  
21 network addresses associated with their respective clients 102 and 104.

22 The client (1) 102 can use the echo server 112 to obtain the network address  
23 associated with the active network device (1) 108. The client (1) can send the echo server  
24 112 a message. When the echo server 112 receives the message from the client (1) 102,  
25



1 the message will include the network address of the active network device 108. In  
2 response, the echo server 112 sends the network address of the active network device 108  
3 back to the client 102. The client (2) 104 may similarly obtain the network address of the  
4 active network device (2) 110. Using the network addresses of the active network devices  
5 108 and 110, the clients 102 and 104 may be able to establish a peer-to-peer connection  
6 124.

7 Although the exemplary environment 100 in Fig. 1 depicts only two clients 102  
8 and 104 in a conversation, it is to be understood that more than two clients may be  
9 involved in a conversation. Two or more clients may communicate in a multipoint  
10 fashion, wherein each client may have a peer-to-peer connection to every other client.  
11 Alternatively, the two or more clients may engage in the conversation through the  
12 switchboard server 114. In addition, any combination of peer-to-peer connections and  
13 connections through the switchboard may be implemented in a conversation. More  
14 detailed descriptions of exemplary operations and systems that may be employed in the  
15 network environment 100 are provided below.  
16

### 17 **Exemplary Systems for Transporting Instant Messages Among Clients**

18 Fig. 2 illustrates a transport system 200 including exemplary layers for managing  
19 communication of data in an instant messaging scenario. In general, each of the layers  
20 provides functionality and/or data for carrying out various tasks associated with the layer.  
21 Describing transport protocols and systems in terms of layers is understood by those  
22 skilled in the art.  
23  
24  
25

1 The transport system 200 includes features 202. Features 202 are functions hosted  
2 or executed by or within a messenger application (e.g., the instant messaging platform  
3 116, Fig. 1) to present data associated with the feature. A feature may be characterized by  
4 the type of data the feature presents, the manner of presenting the data, the operations that  
5 the feature may perform on the data, and/or the interactive options that the feature  
6 provides to the user to interact with the data. For example, a custom user tile feature  
7 presents picture data in a screen tile on the user interface; a file transfer feature enables a  
8 user to select a file and send the file to an instant messaging contact. By way of example,  
9 but not limitation, the features layer 202 may include emoticons, ink, embedded pictures,  
10 and others.

11 The features 202 use layers in a transport protocol stack 206 to communicate data  
12 in an instant messaging scenario. The transport protocol stack 206 transparently manages  
13 details of receiving and transmitting data from and to clients in a conversation. The  
14 transport protocol stack 206 provides a transport application program interface (API) 208  
15 to the features 202, whereby the features 202 can send and receive data. Lower layers of  
16 the transport protocol stack 206 use network configuration data, and information about  
17 the data from the features 202, to establish instant messaging sessions. Thus, the  
18 transport protocol stack 206 makes lower level details of communicating data transparent  
19 to the features 202.  
20

21 The terms "BLOB" (binary large object) and "chunk" are used herein to refer to  
22 portions of data that may be sent and/or received in a network environment, regardless of  
23 the type or size of the data. For example, a digital image may be composed of one or  
24 more BLOBs of data. A BLOB is composed of one or more chunks of data.  
25

1 As mentioned above, the exemplary transport API 208 provides methods, or  
2 functions, that expose functionality of the transport protocol stack to the features 202. A  
3 particular implementation of the API 208 includes the following methods:

4	SendData();	// Sends a blob of data to another client;
	OnDataReceived();	// Callback to feature when associated session receives a blob
		// of data;
5	RegisterFeatureHandler();	// Called by features to register handlers. The handlers are
		// called when another client sends an invitation for the feature.
6	RegisterDataChannelCallbacks();	// Called by features to register handlers that are called when
		// channel events happen or channel states changes.
7	CreateDataChannel();	// Creates a data channel (session) for a feature.
	GetTransportCapabilities();	// Returns capabilities of the conversation, for example Direct,
8		// Indirect, Store and Forward.
	GetParticipants();	// Gets the list of participants in a conversation.

9  
10 To illustrate the exemplary API 208, a feature in the features layer 202 can create  
11 a session using the CreateDataChannel() function. The feature may call the register  
12 RegisterFeatureHandler() to register a handler to be called when an invitation is received  
13 for the feature. The feature may call RegisterDataChannelCallbacks() to register a  
14 handler to be called when a channel event occurs. The feature can call SendData() to send  
15 data to a client. The function OnDataReceived() is called by another transport protocol  
16 layer, such the session layer 210, when data is received for the feature.

17  
18 An exemplary session layer 210 includes logic for providing session services to  
19 features 202. An instant messaging conversation includes at least one session, called an  
20 instant messaging session for carrying on the conversation. An implementation of the  
21 SLP also includes a master session, which has control logic for establishing other sessions  
22 for the features, called feature sessions. The session layer 210 is the protocol by which  
23 the sessions are initiated. In one implementation, the session layer 210 initiates a session  
24 by negotiating session establishment between two peers based on the peers' respective  
25 capabilities.

1 In a particular implementation of the session layer 210, a subset of the session  
2 initiation protocol (SIP) is employed. In this implementation, messages can be sent in a  
3 human-readable format (e.g., Extensible Markup Language (XML), comma separated  
4 value (CSV), etc.). Exemplary session management messages are "INVITE", "BYE",  
5 "CANCEL", and "ACK."

6 An exemplary transport layer 212 includes logic for handling a Transport Layer  
7 Protocol of the Open System Interconnection (OSI) standard. The TLP is a binary  
8 protocol for sending and receiving arbitrary sized binary large objects (BLOBs) via  
9 network channels. The transport layer 212 performs various data management tasks. For  
10 example, as discussed in further detail below, the transport layer 212 is operable to  
11 multiplex data via a network connection to thereby create virtual channels within a single  
12 network connection. As discussed in further detail below, transport layer 212 interfaces  
13 with one or more transport bridges, such as a TCP bridge or an SB-Bridge, which reside  
14 at the transport bridge layer 214. The transport layer 212 may include a scheduler and  
15 queues for scheduling data from the features 202 to be transported via bridges in the  
16 bridge layer 214.  
17

18 In a particular implementation, the transport layer 212 appends a TLP header to  
19 each chunk of each BLOB that is transported. The header provides information about the  
20 BLOB and the chunk. Upon receipt of a chunk, the transport layer 212 can parse the  
21 header to determine how chunks should be reassembled and how BLOBS should be re-  
22 ordered, if necessary. In one exemplary implementation, the TLP header is as follows:

23       DWORD SessId;           // slave session id that this blob belongs to  
24       DWORD BlobId;          // id for this blob  
25       UINT64 Offset;          // offset of this chunk's data into the blob's data  
26       UINT64 BlobSize;        // total size of all chunks not including headers  
27       DWORD ChunkSize;        // size of the data in this chunk (not including header)

DWORD Flags; // ORed together TLP\_CONTROL\_FLAGS

The field “SessId” identifies the session associated with a chunk of data that is being sent or received. The “SessId” field is used by the transport layer 212 to associate a chunk of data with a particular session, and thereby multiplex and/or demultiplex the chunk of data. Data multiplexing employed by the transport protocol stack 206 is described in further detail below.

The field “BlobId” identifies the binary large object (BLOB) corresponding to the data being sent or received. The field “BlobId” may be used by the transport protocol stack 206 to assemble or reassemble chunks into a corresponding BLOB. The “Offset” field represents where a chunk is located in the corresponding BLOB. The BlobId may also be used to organize BLOBs in their proper order before delivery or after receipt. The “BlobSize” field represents the size of the BLOB. The “ChunkSize” field represents the size of the chunk of data, not including the header.

The “Flags” field provides a number of flags used by the transport layer 212 to manage the process of sending and receiving data among clients. Exemplary “Flags” are shown and described as follows:

```
enum TLP_CONTROL_FLAGS
{
    TLP_NAK = 0x01,    // set in a non-acknowledgement packet (request for
                        // retransmit)
    TLP_ACK = 0x02,    // set in an acknowledgement packet
    TLP_RAK = 0x04,    // set in a request for acknowledgement packet
    TLP_RST = 0x08,    // set in a Reset packet, closes the session on the other side
    TLP_BIG = 0x10,    // set when this chunk is part of a blob that should be saved as
                        // a file
    TLP_CAL = 0x20,    // set in a blob for which progress callbacks are requested
};
```

The transport bridge layer 214 includes logic for transporting data over a network. For example, the transport bridge layer 214 may perform data encoding and/or decoding.

1 In one implementation, the transport bridge layer 214 includes one or more transport  
2 bridges that each have functionality corresponding to a network configuration, and data  
3 characteristics, for communicating on the network. Exemplary transport bridges, as well  
4 as systems and operations for selecting among the bridges, are described in detail below.

5 In one implementation, the transport layer 214 can use the transport bridge layer  
6 214 to transition between two types of bridges in the midst of a session. Thus, a session  
7 may begin over a switchboard (SB) bridge, whereby the session data is communicated  
8 through a switchboard server, but, later, the session may switch to a direct, or peer-to-peer  
9 connection. In addition, if the peer-to-peer connection is broken, the transport layer 212  
10 can responsively switch back to the SB bridge or some other available bridge in the  
11 bridge layer 214. The transport layer 212 handles the switch to a different connection  
12 type in a way that is transparent to higher layers of the transport protocol stack 206.  
13

14 The transport layer 212 works with the bridge layer 214 to ensure that BLOBs are  
15 delivered and/or received in proper order. Because the type of connection may change  
16 from a slower connection to a faster connection in the midst of a session, BLOBs that  
17 were transmitted over the slower connection may arrive at their destination after BLOBs  
18 that are transmitted over the faster connection, unless the transport layer 212 takes steps  
19 to ensure the proper order. Exemplary operations are described below that ensure a  
20 proper sequential order of delivery is maintained.

21 A network layer 218 represent lower level communications functionality that  
22 perform standard network communications functions. The network layer 218 may  
23 include, but is not limited to, Transmission Control Protocol/Internet Protocol (TCP/IP)  
24 functionality, Universal Datagram Protocol (UDP) functionality, Hypertext Transport  
25



Protocol (HTTP) functionality, Simple Object Access Protocol (SOAP). The bridge layer 214 can use the network layer 218 functionality to communicate data via a network.

As discussed herein, a number of transport bridges are provided to handle different network configurations, and data types, formats, and/or sizes. Table 1 illustrates exemplary transport bridges and associated network configurations in which the bridges may be employed.

**Table 1**

Bridge	No NAT	1 NAT	2 NATs (not symmetrical)	1 UPnP	1 UPnP & 1 NAT	2 UPnP	Firewall	Multipoint
SB	✓	✓	✓	✓	✓	✓	✓	✓
Twister (IPv6) TCP	✓	✓	✓	✓	✓	✓		
RUDP	✓	✓	✓	✓	✓	✓		

A checkmark (✓) at an intersection of a row and column in Table 1 indicates that the bridge in the intersecting row can be used in the network configuration in the intersecting column.

The SB Bridge sends traffic via a switchboard server (e.g., the switchboard server 114, Fig. 1). In one implementation, when an instant messaging session is initially established, the SB bridge is the bridge used to transport the instant messaging session data. Subsequently, the SB bridge continues to be used only if a direct (peer-to-peer) connection cannot be established.



1       The Twister transport control protocol (TCP) (also called the Internet Protocol  
2       Version 5 (IPv6) TCP) bridge can be used to traverse an Internet Protocol Version 4  
3       (IPv4) NAT. In one implementation, the IPv6 bridge utilizes a Teredo tunneling  
4       technique to establish a peer-to-peer connection. Teredo is an IPv6/IPv4 transition  
5       technology that provides address assignment and host-to-host automatic tunneling for  
6       unicast IPv6 connectivity when IPv6/IPv4 hosts are located behind one or multiple IPv4  
7       NATs. To traverse IPv4 NATs, IPv6 packets are sent as IPv4-based User Datagram  
8       Protocol (UDP) messages.

9       The Twister TCP bridge is operable to establish a peer-to-peer TCP/IP  
10      connection, even if a network address translator (NAT) is configured on the network. For  
11      example, if a first client is behind a NAT, and a second client is not behind a NAT or  
12      firewall, the twister TCP bridge of the first client initiates an outgoing TCP connection  
13      with the second client. The first client can use an echo server (e.g., the echo server 112,  
14      Fig. 1) to detect that the first client is behind the NAT, and thus determine the direction in  
15      which to form the connection; i.e., the first client initiates the TCP connection. A similar  
16      technique may be applied when one of the clients is behind a Universal Plug and Play  
17      (UPnP) NAT.

18      The Reliable Universal Datagram Protocol (RUDP) bridge communicates data via  
19      UDP but emulates a TCP connection by acknowledging receipt of transmission of data  
20      much as a "packet" would be acknowledged under TCP. In this fashion, the RUDP bridge  
21      provides a more reliable than a traditional UDP connection, because the protocol can  
22      detect when data is dropped (i.e., not received) and should be resent. The RUDP bridge  
23      may use an echo server to obtain NAT network addresses, so that NATs can be traversed.  
24  
25

1           The RUDP bridge appends a header to transmitted data. An exemplary header is  
2 shown below:

```
3           unsigned __int64 qwSeqNum;        // sequence number of first byte in this packet  
4           unsigned __int64 qwAckNum;        // ack number of next byte the sender of this packet expects  
5           DWORD dwOtherFields;            // 4 bit offset in quadwords, 5 bit reserved, 5 control bits, 16  
6                                            // reserved, from most significant  
7           DWORD dwWindow;                 // window size receiver will accept  
8           DWORD dwSendTime;                // timestamp for when the packet was sent  
9           DWORD dwEchoTime;                // last timestamp the sender recorded in a packet from receiver
```

10           An exemplary RUDP bridge protocol involves sending a “SYN” message to a  
11 listening RUDP bridge, responding with “SYN ACK”, and responding to the “SYN  
12 ACK” message with an “ACK.” Subsequently, a retransmit timer will continue to  
13 retransmit one or more sets of data until an ACK is received from the responding RUDP  
14 bridge for the transmitted (or retransmitted) data sets.

15           Fig. 3 is a block diagram illustrating an exemplary selection module 302 operable  
16 to select a preferred transport bridge from among a set of transport bridges 304. The  
17 instant messaging application may have more than one feature 308, and each feature 308  
18 communicates with a master session 306 and a session registrar 310 to set up an  
19 associated feature session 312. Instant messaging data, including feature data, is  
20 transported through a selected bridge.

21           The master session 306 executes at the session layer 210 (Fig. 2) in the transport  
22 protocol stack 206 (Fig. 2). The master session 306 is created when a conversation is  
23 started in an instant messaging application. The master session 306 is associated with a  
24 conversation window and is the object through which features 308 request slave, or  
25 feature sessions 312. The master session 306 may also expose a static method that

1 enables features 308 to start a master session outside the context of a conversation  
2 window, whereby an instant messaging conversation will be created.

3 A feature 308 registers with a session registrar 310, which maintains information  
4 about features running during a conversation. The session registrar 310 may also  
5 maintain information about “auto-start” features. Auto-start features are features that  
6 should exist by default and their invite handlers will be invoked at both clients in the  
7 conversation with predefined call IDs and session IDs once an underlying switchboard  
8 connection is established.

9 When the feature 308 registers with the session registrar 310, the feature 308  
10 provides an invitation handler function and a globally unique identifier (GUID)  
11 associated with the feature 308. The master session 306 can use the feature registrar  
12 information to establish the feature session 312, through which the feature 308 can carry  
13 on a session with a peer feature.

14 For example, an invitation from a peer may be received by the master session 306,  
15 inviting the feature 308 to enter into a session. The invitation specifies the GUID  
16 associated with the feature 308. The master session 306 looks up the GUID in the session  
17 registrar 310, to find the invitation handler function associated with the feature 308. If  
18 the invitation is accepted by the feature 308, the master session 306 creates the new  
19 feature session 312.  
20

21 The feature 308 may invite a peer feature to enter into a session. To do so, the  
22 feature 308 requests a new session from the master session 306 (for example, by calling  
23 CreateDataChannel() in the transport API 208, Fig. 2)). The request for a new session  
24 from the feature 308 can include a request for a type of connection (e.g., direct, non-  
25

1 direct). The master session 306 will call the bridge selector 302 with information such as  
2 the direct connection requested, the size of the data, and/or the network configuration.  
3 The bridge selector 302 uses the information to identify a preferred transport bridge to use  
4 for the requested feature session.

5 In one implementation, the bridge selector 302 executes at the transport layer 212  
6 (Fig. 2) and implements transport bridge selection rules to select the preferred bridge. In  
7 one implementation, the design goal is to send as much traffic directly peer-to-peer as  
8 possible, thereby avoiding delays and costs that may be associated with a switchboard  
9 server. In this implementation, the following factors are analyzed to select the transport  
10 bridge: the direct connection requested, the bandwidth supported by the transport bridge,  
11 the size of data to be sent, any limitations of the transport bridge, the cost associated with  
12 a particular transport bridge (e.g., relay costs), and/or privacy specifications (e.g., reveal  
13 or not reveal IP address).

14 Each transport bridge computes a score based on the above factors. The bridge  
15 gives its score to the bridge selector 302, which chooses the preferred transport bridge  
16 based on the scores. Before analyzing the raw scores, the bridge selector 302 takes  
17 specified requests into account. For example, if the feature 308 requests a direct  
18 connection, then the bridge selector 302 may eliminate all non-direct bridges from the set  
19 of candidate bridges. Bridges that will not work in a particular network configuration are  
20 also eliminated from the set of candidate bridges. For example, the Twister TCP bridge  
21 is not considered if the network configuration includes 2 symmetric legacy NATs.  
22 Information obtained from an echo server test during conversation initialization can be  
23 used to decide which bridges cannot work.  
24  
25

1 Thus, one implementation of the bridge selector 302 employs the following  
2 transport bridge selection rules:

- 3 1. Eliminate bridges that cannot work based on echo test results.
- 4 2. Eliminate bridges that won't satisfy specific requests by caller.
- 5 3. Use remaining bridge with highest score.
- 6 4. If previous bridge fails, use remaining bridge with second highest score.
- 7 5. If previous bridge fails, use remaining bridge with next highest, and so on...

8 The bridge selector 302 notifies the master session 306 of the preferred bridge  
9 based on the rules. In response, the master session 306 creates the feature session 312  
10 and instantiates the preferred bridge to be used. After the feature session 312 is created,  
11 the feature session 312 will be the session through which the feature 308 communicates  
12 with other peer features. The feature session 312 handles sending data, as well as session  
13 control calls, such as AcceptSession and AcceptDirect.

14 Fig. 4 is a block diagram illustrating virtual channels 402 across a physical  
15 connection 404 between a first client device 406 and a second client device 408. The  
16 connection 404 and the channels 402 are used by the client devices 406 and 408 to engage  
17 in an instant messaging conversation, wherein data of various types is communicated.  
18 Although the devices 406 and 408 are referred to as "client devices," it is to be  
19 understood that the connection 404 between the devices 406 and 408 need not be through  
20 a server, but could be a direct, or peer-to-peer connection.

21 As shown, data channel 0 is used to communicate text messages in the  
22 conversation; data channel 1 is used to transfer files between the client devices 406 and  
23 408; data channel 2 is used to transfer embedded pictures between the devices 406 and  
24 408; and data channel 3 is used to communicate typing indicator information that  
25 indicates whether a user is currently typing a message at one of the devices 406 or 408.

1       The virtual channel number corresponds to a session identifier. As discussed  
2 above, a header (see the TLP header above) is appended to data communicated over the  
3 connection 404. The header includes a "SessId" which indicates which session, and  
4 which channel, (e.g., feature or master) the data corresponds to. Using the SessId, data  
5 from multiple sessions can be multiplexed over the connection 404. Thus, the channels  
6 402 are considered "virtual" channels. The SessId is stripped off received data at the  
7 client devices 406 and 408 and the data is routed to the appropriate session associated  
8 with the SessId.

9       In another implementation, more than one connection is established between the  
10 client device 406 and the client device 408, and each of the connections can include  
11 multiple virtual channels. For example, when a conversation is initially created between  
12 the client device 406 and the client device 408, a master session will be created via a  
13 connection through a switchboard server. Later, a peer-to-peer connection may be  
14 established between the client devices 406 and 408. In this situation, features on the  
15 devices 406 and 408 may establish virtual channels on either the switchboard connection  
16 or the peer-to-peer connection.  
17

### 18 19 **Exemplary Operations Employing a Transport Protocol Stack**

20       Fig. 5 illustrates an operation flow 500 having exemplary operations for  
21 establishing a feature session within a master session and creating a direct connection if  
22 possible. The operation flow 500 may be carried out by a client device, such as the client  
23 computer (1) 102, shown in Fig. 1, in order to engage in an instant messaging  
24 conversation with another client device.  
25



1           An establishing operation 502 establishes a master session. The master session is  
2 established when a user initially launches the messenger application and attempts to  
3 contact another user. Establishing the master session involves inviting another client to  
4 enter into the conversation. After the other client accepts the invitation, the master  
5 sessions are created on both the inviting client and the accepting client. The master  
6 session is established through a switchboard bridge, via an SB connection. The master  
7 session may be used later to establish feature sessions.

8           If an echo server test has not been conducted prior to the establishing operation  
9 502, the establishing operation 502 may conduct an echo server test, to detect firewalls,  
10 NATs, or other active network devices that may be present. The echo server test involves  
11 sending a test message to an echo server (e.g., the echo server 112, Fig. 1), and the echo  
12 server responding with network address information corresponding to any active network  
13 devices that may be configured.

14           The user may want to communicate feature data in addition to the text data. Such  
15 feature data includes, but is not limited to files, pictures, custom emoticons, custom user  
16 tiles, and the like. In addition, the messenger application may create a feature session to  
17 transfer feature data. Thus, in another establishing operation 504, a feature session is  
18 established. The establishing operation 504 involves the master session creating a feature  
19 session at an inviting client and sending a feature invitation to a receiving client. When  
20 the feature invitation arrives at the receiving client, the invitation is either accepted or  
21 rejected.  
22

23           Continuing with the establishing operation 504, if the feature invitation is  
24 accepted, the master session on the receiving client creates a feature session on the  
25



1 receiving client. The feature session on the receiving client has the GUID corresponding  
2 to the inviting feature. The master session on the receiving client also calls an invitation  
3 handler function that is registered for the GUID. If the feature invitation is accepted, the  
4 associated feature stores the feature session, so that the feature session can be used to  
5 interact with the transport API for the session. When the feature accepts, callback  
6 messages are sent to both the inviting feature and the accepting feature letting them know  
7 the session is ready for use to send and receive data. If the feature invitation is rejected,  
8 the master session deletes the feature session that was created.

9 The establishing operation 504 establishes the feature session through an existing  
10 connection, such as an established SB connection. More particularly, the feature session  
11 data is multiplexed with the master session data, as shown in Fig. 4, to carry on the  
12 feature session and the master session simultaneously. Because a connection through the  
13 switchboard may be slower than a peer-to-peer connection, a creating operation 506  
14 attempts to create a direct (peer-to-peer) connection over which to conduct the feature  
15 session. An exemplary embodiment of the creating operation 506 is illustrated in Fig. 6,  
16 and described below. If a direct connection can be created, the feature session will be  
17 carried out using the direct connection. The operation flow 500 ends at an ending  
18 operation 508.  
19

20 Fig. 6 illustrates an operation flow having exemplary operations for creating a  
21 direct connection. It is assumed that a user indicates to a feature at an initiating client that  
22 the user wants to change the feature session to a direct connection rather than the  
23 switchboard (SB) server. The user may prefer the direct connection because she wants to  
24 avoid being limited to a message size limit, or she wants to send more messages than  
25

1 allowed by the SB server, or perhaps she wishes to keep her messages inside her  
2 corporate firewall.

3 When the feature receives the indication to move to a direct connection, a  
4 requesting operation 602 executes, wherein the feature calls a method on the feature's  
5 slave session. The requesting operation 602 requests a direct connection. The slave  
6 session passes the request to the master session. In response, the master session sends a  
7 "reINVITE" message to the peer feature on a receiving client requesting a direct  
8 connection.

9 Continuing with the requesting operation 602, the "reINVITE" message contains  
10 information that identifies all transport bridge types supported by the inviting client. The  
11 feature session on the receiving client calls the invite handler function associated with the  
12 feature. The "reINVITE" may be accepted in any number of ways. In one  
13 implementation, the feature may have set an auto-accept setting in the feature's invitation  
14 handler. An auto-accept setting indicates that all requests for a direct connection are to be  
15 automatically accepted. In another implementation, an auto-reject setting may be set in  
16 the invite handler, which indicates that the direct connection request will be automatically  
17 rejected.  
18

19 If neither an auto-accept or auto-reject setting is set, the transport will look to a  
20 global default value. If the global default value is set to accept or reject, the appropriate  
21 reply is made to the "reINVITE." If it is not set, the invited feature can respond to the  
22 invitation, for example, by prompting the user as to whether to accept or reject the  
23 "reINVITE." Assuming the "reINVITE" is accepted, a preferred transport bridge is  
24 identified in a selecting operation 606.  
25

1       The selecting operation 606 selects a preferred transport bridge using preference  
2 factors as discussed above with respect to Fig. 3. An accepting operation 608 sends an  
3 accept message back to the inviting client. The accept message contains the preferred  
4 bridge. The acceptance also includes the accepting client's internal and external IPv4  
5 addresses as well as an IPv6 global address if the client has one.

6       The transport on the other end will attempt to connect to the received IP address  
7 and port using the preferred bridge listed in the accept message. After the inviting client  
8 has connected, the transport sends a "RAK" and the accepting client replies with an  
9 "ACK." If the connection fails, the inviting client will time out after sending the "RAK".  
10 In one implementation, the inviting master session then sends another "reINVITE," this  
11 time excluding the transport bridge type that failed. The operation flow ends at an ending  
12 operation 610.

13       Fig. 7 illustrates an operation flow 700 having exemplary operations for  
14 transporting binary large objects (BLOBs) in sequential order. In general, before a last  
15 chunk of a BLOB is transmitted over a new direct connection, the transport waits to  
16 receive all "ACKs" associated with chunks of BLOBs that were previously transmitted  
17 over a non-direct connection. Waiting for "ACKS" in this fashion occurs, regardless of  
18 what type of transport bridge or connection (i.e., UDP, TCP, or otherwise) is being used.

19       A transmitting operation 702 begins transmitting a BLOB using a non-direct  
20 connection. A query operation 704 determines whether the transport bridge type has been  
21 switched to a direct connection. If the transport bridge type has not been switched the  
22 query operation 704 branches "NO" to a transmitting operation 706. The transmitting  
23 operation 706 transmits a chunk of the BLOB via the non-direct connection.  
24  
25

1 After the chunk is transmitted in the transmitting operation 706, the query  
2 operation 704 again determines whether the transport bridge type has changed. If the  
3 transport bridge type has changed, the query operation 704 branches "YES" to another  
4 query operation 708. The query operation 708 determines if a current BLOB identifier  
5 (ID) is greater than the last BLOB ID. If the current BLOB ID is not greater than the last  
6 BLOB ID, the query operation 708 branches "NO" to the transmitting operation 706,  
7 which transmits a chunk of the current BLOB.

8 If the query operation 708 determines that the current BLOB ID is greater than the  
9 last BLOB ID, then the query operation 708 branches "YES" to another query operation  
10 710. The query operation 710 determines whether the current chunk to be transmitted is  
11 the last chunk in the BLOB. If the current chunk is not the last chunk in the BLOB, the  
12 query operation 710 branches "NO" to the transmitting operation 706, which transmits  
13 the current chunk. If the query operation 710 determines that the current chunk is the last  
14 chunk in the BLOB, the query operation 710 branches to a waiting operation 712. The  
15 waiting operation 712 waits until all ACKs are received for BLOBs that were transmitted  
16 via the original non-direct connection.  
17

18  
19 **An Exemplary Computer System that may be used in a Network Environment for**  
20 **Instant Messaging**

21 Fig. 8 and the corresponding discussion are intended to provide a general  
22 description of a suitable computing environment in which the described arrangements and  
23 procedures for transporting computer data may be implemented. Exemplary computing  
24 environment 820 is only one example of a suitable computing environment and is not  
25

1 intended to suggest any limitation as to the scope of use or functionality of the described  
2 subject matter. Neither should the computing environment 820 be interpreted as having  
3 any dependency or requirement relating to any one or combination of components  
4 illustrated in the exemplary computing environment 820.

5 The exemplary arrangements and procedures to transport computer data between  
6 interconnected devices are operational with numerous other general purpose or special  
7 purpose computing system environments or configurations. Examples of well known  
8 computing systems, environments, and/or configurations that may be suitable for use with  
9 the described subject matter include, but are not limited to, personal computers, server  
10 computers, thin clients, thick clients, hand-held or laptop devices, multiprocessor  
11 systems, microprocessor-based systems, mainframe computers, distributed computing  
12 environments such as server farms and corporate intranets, and the like, that include any  
13 of the above systems or devices.

14 The computing environment 820 includes a general-purpose computing device in  
15 the form of a computer 830. The computer 830 may include and/or serve as an  
16 exemplary implementation of a transport system for instant messaging described above  
17 with reference to Figs. 1-7. The components of the computer 830 may include, by are not  
18 limited to, one or more processors or processing units 832, a system memory 834, and a  
19 bus 836 that couples various system components including the system memory 834 to the  
20 processor 832.

21 The bus 836 represents one or more of any of several types of bus structures,  
22 including a memory bus or memory controller, a peripheral bus, an accelerated graphics  
23 port, and a processor or local bus using any of a variety of bus architectures. By way of  
24  
25

1 example, and not limitation, such architectures include Industry Standard Architecture  
2 (ISA) bus, Micro Channel Architecture (MCA) bus, Enhanced ISA (EISA) bus, Video  
3 Electronics Standards Association (VESA) local bus, and Peripheral Component  
4 Interconnects (PCI) bus also known as Mezzanine bus.

5 The computer 830 typically includes a variety of computer readable media. Such  
6 media may be any available media that is accessible by the computer 830, and it includes  
7 both volatile and non-volatile media, removable and non-removable media.

8 The system memory includes computer readable media in the form of volatile  
9 memory, such as random access memory (RAM) 840, and/or non-volatile memory, such  
10 as read only memory (ROM) 838. A basic input/output system (BIOS) 842, containing  
11 the basic routines that help to communicate information between elements within the  
12 computer 830, such as during start-up, is stored in ROM 838. The RAM 840 typically  
13 contains data and/or program modules that are immediately accessible to and/or presently  
14 be operated on by the processor 832.

15 The computer 830 may further include other removable/non-removable,  
16 volatile/non-volatile computer storage media. By way of example only, Fig. 8 illustrates  
17 a hard disk drive 844 for reading from and writing to a non-removable, non-volatile  
18 magnetic media (not shown and typically called a "hard drive"), a magnetic disk drive  
19 846 for reading from and writing to a removable, non-volatile magnetic disk 848 (e.g., a  
20 "floppy disk"), and an optical disk drive 850 for reading from or writing to a removable,  
21 non-volatile optical disk 852 such as a CD-ROM, DVD-ROM or other optical media. The  
22 hard disk drive 844, magnetic disk drive 846, and optical disk drive 850 are each  
23 connected to bus 836 by one or more interfaces 854.  
24  
25



1       The drives and their associated computer-readable media provide nonvolatile  
2 storage of computer readable instructions, data structures, program modules, and other  
3 data for the computer 830. Although the exemplary environment described herein  
4 employs a hard disk, a removable magnetic disk 848 and a removable optical disk 852, it  
5 should be appreciated by those skilled in the art that other types of computer readable  
6 media which can store data that is accessible by a computer, such as magnetic cassettes,  
7 flash memory cards, digital video disks, random access memories (RAMs), read only  
8 memories (ROM), and the like, may also be used in the exemplary operating  
9 environment.

10       A number of program modules may be stored on the hard disk, magnetic disk 848,  
11 optical disk 852, ROM 838, or RAM 540, including, by way of example, and not  
12 limitation, an operating system 858, one or more application programs 860, other  
13 program modules 862, and program data 864. Application programs 860 may include an  
14 instant messaging application, utilizing features and a transport protocol stack, as  
15 discussed herein.  
16

17       A user may enter commands and information into the computer 830 through  
18 optional input devices such as a keyboard 866 and a pointing device 868 (such as a  
19 “mouse”). Other input devices (not shown) may include a microphone, joystick, game  
20 pad, satellite dish, serial port, scanner, or the like. These and other input devices are  
21 connected to the processing unit 832 through a user input interface 870 that is coupled to  
22 the bus 836, but may be connected by other interface and bus structures, such as a parallel  
23 port, game port, or a universal serial bus (USB).  
24  
25



1 An optional monitor 872 or other type of display device is connected to the  
2 bus 836 via an interface, such as a video adapter 874. In addition to the monitor, personal  
3 computers typically include other peripheral output devices (not shown), such as speakers  
4 and printers, which may be connected through output peripheral interface 875.

5 The computer 830 may operate in a networked environment using logical  
6 connections to one or more remote computers, such as a remote computer 882. The  
7 remote computer 882 may include many or all of the elements and features described  
8 herein relative to the computer 830. The logical connections shown in Fig. 8 are a local  
9 area network (LAN) 877 and a general wide area network (WAN) 879. The LAN 877  
10 and/or the WAN 879 can be wired networks, wireless networks, or any combination of  
11 wired or wireless networks. Such networking environments are commonplace in offices,  
12 enterprise-wide computer networks, intranets, and the Internet.

13 When used in a LAN networking environment, the computer 830 is connected to  
14 the LAN 877 via a network interface or an adapter 886. When used in a WAN  
15 networking environment, the computer 830 typically includes a modem 878 or other  
16 means for establishing communications over the WAN 879. The modem 878, which may  
17 be internal or external, may be connected to the system bus 836 via the user input  
18 interface 870 or other appropriate mechanism. Depicted in Fig. 8 is a specific  
19 implementation of a WAN via the Internet. The computer 830 typically includes a modem  
20 878 or other means for establishing communications over the Internet 880. The modem  
21 878 is connected to the bus 836 via the interface 870.  
22

23 In a networked environment, program modules depicted relative to the personal  
24 computer 830, or portions thereof, may be stored in a remote memory storage device. By  
25

1 way of example, and not limitation, Fig. 8 illustrates remote application programs 889 as  
2 residing on a memory device of remote computer 882. It will be appreciated that the  
3 network connections shown and described are exemplary and other means of establishing  
4 a communications link between the computers may be used.

## 5 6 **Conclusion**

7 Although the described arrangements and procedures to have been described in  
8 language specific to structural features and/or methodological operations, it is to be  
9 understood that the subject matter defined in the appended claims is not necessarily  
10 limited to the specific features or operations described. Rather, the specific features and  
11 operations are disclosed as preferred forms of implementing the claimed present subject  
12 matter.  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25